

WestminsterResearch

<http://www.westminster.ac.uk/westminsterresearch>

**User experience evaluation of human representation in
collaborative virtual environments**

Economou, D., Doumanis, I., Argyriou, L. and Georgalas, N.

This is the published version of Economou, D., Doumanis, I., Argyriou, L. and Georgalas, N. (2017) User experience evaluation of human representation in collaborative virtual environments., *Personal and Ubiquitous Computing*, DOI: [10.1007/s00779-017-1075-4](https://doi.org/10.1007/s00779-017-1075-4). It is available from the publisher at:

<https://doi.org/10.1007/s00779-017-1075-4>

© The Author(s) 2017. This article is published with open access at Springerlink.com

The WestminsterResearch online digital archive at the University of Westminster aims to make the research output of the University available to a wider audience. Copyright and Moral Rights remain with the authors and/or copyright owners.

Whilst further distribution of specific materials from within this archive is forbidden, you may freely distribute the URL of WestminsterResearch: (<http://westminsterresearch.wmin.ac.uk/>).

In case of abuse or copyright appearing without permission e-mail repository@westminster.ac.uk

User experience evaluation of human representation in collaborative virtual environments

Daphne Economou¹ · Ioannis Doumanis² · Lemonia Argyriou¹ · Nektarios Georgalas³

© The Author(s) 2017. This article is an open access publication

Abstract Human embodiment/representation in virtual environments (VEs) similarly to the human body in real life is endowed with multimodal input/output capabilities that convey multiform messages enabling communication, interaction and collaboration in VEs. This paper assesses how effectively different types of virtual human (VH) artefacts enable smooth communication and interaction in VEs. With special focus on the REal and Virtual Engagement In Realistic Immersive Environments (REVERIE) multi-modal immersive system prototype, a research project funded by the European Commission Seventh Framework Programme (FP7/2007-2013), the paper evaluates the effectiveness of REVERIE VH representation on the foregoing issues based on two specifically designed use cases and through the lens of a set of design guidelines generated by previous extensive empirical user-centred research. The impact of REVERIE VH representations on the quality of user experience (UX) is evaluated through field trials. The output of the current study proposes directions for improving human

representation in collaborative virtual environments (CVEs) as an extrapolation of lessons learned by the evaluation of REVERIE VH representation.

Keywords Virtual environments · Virtual humans · Avatars · Design guidelines · Immersion · Presence · Communication · Interaction · Collaboration

1 Introduction

Human embodiment/representation within collaborative virtual environments (CVEs), to which we will refer in the rest of the paper as virtual humans (VHs), is very important if one considers the role of the human body in real life. The human body

- Provides immediate and continuous information about user presence, identity, attention [1], activity status, availability [2] and mood
- Sets a social distance between conversants (e.g. their actual location) [3] which helps in regulating interaction [4]
- Helps managing a smooth sequential exchange between parties by supporting speech with non-verbal communication [5]

Non-verbal communication or bodily communication “takes place whenever one person influences another by means of facial expression, tone of voice, or any other channels (except linguistic)” [6] (involving body language, hand gesture, gaze and facial expressions, or any combination thereof). VHs in VEs play the same role as the human body does in real life, and thus, they have been recognized as key elements for human interaction and communication in CVEs [7, 8]. Another important function of VHs is to help

✉ Daphne Economou
D.Economou@westminster.ac.uk

Ioannis Doumanis
ioannis@ctvc.co.uk

Lemonia Argyriou
argyrioulemonia@gmail.com

Nektarios Georgalas
Nektarios.georgalas@bt.com

¹ Department of Computer Science, Faculty of Science and Technology, University of Westminster, London, UK

² CTVC Ltd, London, UK

³ BT Intel Co-lab, British Telecom, Ipswich, UK

humans immerse themselves in the mediated environment. Immersion in this context refers to the degree the VHs help to create a sensation of being spatially located in the CVE facilitating the users feel as being actually “there” and participate in the virtual world. This feeling of spatial presence has been studied extensively in psychology with about theories developed as a result [9].

Current virtual reality (VR) systems and frameworks enable various types of human representation. However, VHs are still far from fully encompassing the required attributes in order to realistically represent humans in VR and support immersion. In this paper, we review the effectiveness of VHs supported by REal and Virtual Engagement In Realistic Immersive Environments (REVERIE), a VR system prototype developed under the European Community’s FP7 [10] in terms of addressing issues related to virtual presence, communication and interaction in CVEs.

Empirical user-centred research discussed further in Section 5 led to the creation of a list of design guidelines for VHs in CVEs [11, 12], shown in the first and second columns of Table 1. The application of those design guidelines in the design of human embodiment in CVEs ensures smooth remote communication, interaction and collaboration. Those design guidelines are used to evaluate how effectively the different types of REVERIE VHs support smooth communication, interaction and collaborations in CVEs. Design guidelines provide a direct mean of translating detailed design specifications into actual implementation and a way for designers to determine the consequences of their design decisions. They are also a great tool that can drive technological challenges in suggesting directions in which the underlying VHs and CVE technology should be developed to support designers implementing their decisions. Despite the study being conducted some time ago, the design guidelines are still relevant and valid to current CVEs.

The rest of the paper is structured as follows. Section 2 reviews state-of-the-art VR systems in terms of realistic human representation supported in VR. Section 3 is an introduction to the REVERIE project demonstrating different types of VHs supported. Section 4 presents the user scenarios and field trials based on which REVERIE VH representations are being evaluated. Those field trial results are used in Section 5 to compare how well REVERIE VHs address the aforementioned set of VH design guidelines related to virtual presence in CVEs. Section 6 provides a discussion of the importance of VH representation in CVEs. Finally, Section 7 closes with conclusions and future directions.

2 State of the art of virtual human representation

Until recently, most of VH representation platforms enabled human representation within CVEs as animated 3D avatars.

An avatar is a 3D graphical representation of a user’s character that can take any form (cartoon-like, animal-like or anthropomorphic). In addition, avatars in most CVEs are controlled by the keyboard and mouse limiting user actions and interactivity.

Second Life [13], OpenSim [14] and Active Worlds [15] are the first online VR platforms that allowed users to create their own avatars and use them to explore the virtual world or interact with other avatars, places or objects. By means of these avatars, users can meet other VR residents, socialize, participate in individual and group activities, build, create, shop and trade virtual property and services with one another.

Sansar [16] is a VR platform created by Linden Labs as the official successor of Second Life. Sansar aims to democratize VR content creation (including avatars) by empowering people to easily create, share and monetize their content without requiring engineering resources. Although the platform has not been officially released yet, it already features several user-generated worlds of impressive beauty and detail. However, in terms of user representation, Sansar seems to follow the same approach as in Second Life. Users have access to a library of female and male avatars which they can fully customize (e.g., change their skin colour, hair). Using a custom avatar, they can explore VR worlds and communicate with other users by either text or voice. Apart from supporting the latest Oculus Rift headset, Sansar does not (at least in its current version) support any other multimodal technologies (e.g., facial expression mapping or full body avatar puppeting) to increase immersion in the VR environments.

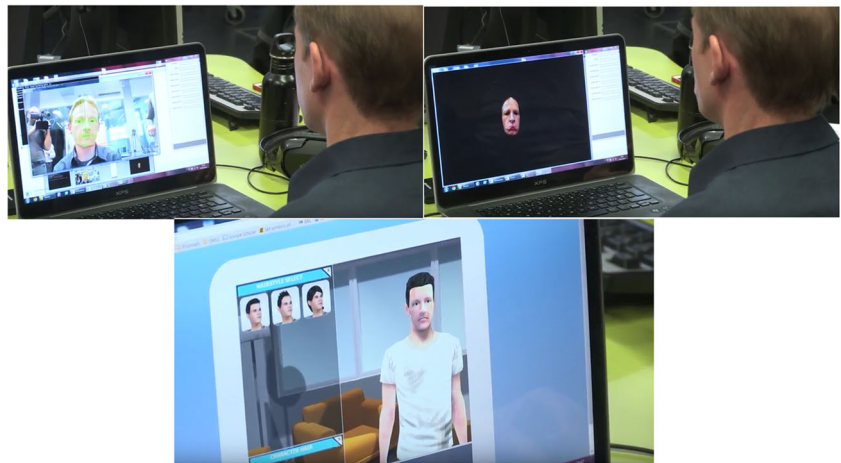
High Fidelity [17] aims to improve human representation and realistic interaction in virtual worlds compared to all of the above VR platforms by using sophisticated motion capture techniques that could mirror user body and head movement, plus facial expressions onto their avatar and allow controlling the avatar arms and torso in order to interact as naturally as possible in the virtual world. The platform supports a range of devices and inputs for greater immersion and control (e.g. Oculus Rift, Leap Motion controller and Microsoft Kinect). Although High Fidelity is a promising virtual reality platform, it does not support the latest generation of VH representation called “replicants”, which refers to a dynamic full-body 3D reconstruction of the user. Replication technology uses the latest generation of motion capture sensors (e.g. Microsoft Kinect) to create a photorealistic representation of the human user in the CVE.

All previous VR platforms do not track or represent the user’s body in real scale with natural motion to the user, due to a lack of data about its position and orientation in the world. The result is that the user body is not visibly a part of the environment, which risks damaging the user’s immersion. SomaVR is a platform that performs an in-depth

Table 1 Mapping of how REVERIE VHS address user-centred VH design guidelines (DGs) that support communication, interaction and collaboration in CVs resulting in enhancing immersion (shows that the DG has not been met, ✓ shows that the DG has been met)

Design guidelines (DGs)	Avatars	Puppeted avatars	Human replicants	ECA
DG1: VHS should support realistic or aesthetically pleasing representation of the user	–	✓	–	–
DG2: VHS should support unique representation	✓	✓	✓	–
DG3: VHS should convey the user's role in the CVE (e.g. student, teacher, other)	✓	✓	✓	✓
DG4: VHS should support customizable behaviour	–	–	–	–
DG5: VHS should convey the user's viewpoint	✓	✓	✓	✓
DG6: User viewpoints should be easily directed to see an active participant or a speaker even when they are out of other user viewpoint	–	–	–	✓
DG7: An active participant needs to be identified even when their VH is out of other users' viewpoints	–	–	–	✓
DG8: A tool should be provided for users to lock onto the active VH and follow it automatically	–	–	–	✓
DG9: A VH should be easily associated with its communication	–	✓	✓	✓
DG10: The speaker needs to be identified even when their VH is out of other users' viewpoints	–	–	–	✓
DG11: VHS should convey the user's intention to take turn or offering a turn even when not being in other users' viewpoints	–	–	–	–
DG12: Private communication and interaction should be supported	–	–	–	–
DG13: VHS should show when the user is involved in private communication and whether or not others could join in	–	✓	✓	–
DG14: VHS should reveal the user's action point	–	✓	✓	–
DG15: Users need to be provided with real-time cues about their own actions	–	✓	✓	–
DG16: VHS should convey explicitly the user's process of activity and state of mind	–	–	✓	✓
DG17: The expert should be in control of novice user behaviour	–	–	✓	✓
DG18: The expert should have control over an individual user's viewpoint	–	–	–	✓
DG19: The expert should be in control of the communication tools	✓	✓	✓	✓
DG20: the expert should be able to take control of objects in the CVE	–	–	–	–
DG21: The expert should be aware of and have control over private communication of novice users	–	–	–	–
DG22: The expert should be aware of and have control over private interactions of novice users	–	–	–	–
DG23: The expert should have an episodic memory of novice user mistakes	–	–	–	–

Fig. 1 RAAT tool preliminary testing



analysis of the data provided by both a Kinect V2 and HTC Vive and creates a virtual body for the user that moves and acts as their own and can be perceived from a natural first-person perspective [18]. SomaVR aims to enable players feel physically grounded as their virtual body replicates their own. However, SomaVR does not support facial expressions.

3 REVERIE technologies for online human representation and interaction

REVERIE is a multimodal and multimedia system prototype that offers a wide range of VH representations and input control mechanisms. The REVERIE framework enables users to meet and share experiences in an immersive VE by using cutting-edge technologies for 3D data acquisition and processing, networking and real-time rendering.

Similarly to the platforms discussed above, REVERIE's standard VH representation is an avatar. However, the platform also enables humans to be represented in the virtual world by real-time realistic full-body 3D reconstructions enabling the creation of more immersive experiences [19]. The system supports also affectively user representation in the virtual worlds by employing various modules that analyse user behaviour in the real world. There are modules to analyse: body gestures, facial expressions, and speech and modules to analyse emotional and social aspects.

Finally, REVERIE's virtual worlds can be cohabited by humans (in the form of avatars or replicants) as well as autonomous agents. Thus, REVERIE is considered as a system that provides currently a more holistic approach in terms of human representation in VR. The different VH representations that REVERIE supports are presented below.

REVERIE VR system prototype features both human-to-human and human-to-agent interaction. Users can enter the VE represented by conventional avatars, puppeted avatars or real-time dynamically reconstructed users (replicants). They can adapt the basic features of VHs, but they can also create photorealistic 3D representations of themselves to interact with other users and embodied conversational agents (ECAs) in VEs.

3.1 Avatars

REVERIE supports an avatar authoring tool (RAAT) [20] that allows the creation of bespoke VHs that closely match the facial appearance of the representative user, see Fig. 1. Users are provided with the option to use RAAT to create their own lookalike personalized avatar simply by allowing the tool to take a single snapshot of their face by using their device's webcam. This personalized lookalike avatar that resembles the users' appearance is their VH representation for navigating and interacting with others in the virtual scenes offered by REVERIE.

3.2 Puppeted avatars

Puppeteering or avateering VHs refers to the process of mapping a user's natural motion and live performance to a VH's deforming control elements in order to realistically reproduce the activity during rendering cycles [21]. REVERIE supports puppeted avatars adopting two different types of technology:

- Kinect Body-Puppeted avatars using a single-sensor device to implement advanced algorithmic solutions which enable user activity analysis, full body reconstruction, avatar puppeting and scene navigation, see Fig. 2.

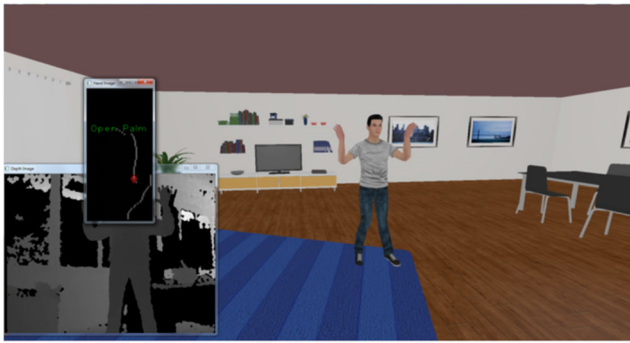


Fig. 2 Kinect Body-Puppeted avatar shown in the 3D Hangout environment using the Kinect gesture-driven navigation module via Kinect skeleton resource sharing offered by the Shared Skeleton common module

- Webcam Face-Puppeted avatars, which is based on modules that perform facial detection, tracking the features by point extraction from a single, front-facing web camera connected to the system while deformation and rendering of the character's face mesh geometry happens in a separate component

In the first option, the user's virtual body representation is controlled by skeleton-based tracking with the use of a Kinect depth sensing camera. The virtual body moves in the virtual scene according to the user movements in the real world. The second option uses a simple web camera to capture the user facial expressions which are then mapped to its avatar's face, replicating in this way user emotions to its virtual representation. Both options are integrated in the REVERIE platform allowing users to select the most suitable one to various use cases. In combination with the RAAT tool, these options offer an efficient way of realistic human representation in VEs.

3.3 Human replicants

The most realistic human representation is achieved in REVERIE by means of "replicants". By replicants, we refer to a dynamic full-body 3D reconstruction of the user. The REVERIE visual "Capturing" module is responsible for capturing a user during a live session with the use of multiple depth-sensing devices (Kinects) and dynamically reconstructing a 3D representation of the user (including both 3D geometry and texture) in real time. The user moves inside a restricted area, surrounded by at least three Kinect devices, while standing in front of the display interacting with other participants. The "replicant" reconstruction is coded in real time and transmitted in order to be visualized on other users' displays along with the elements of the shared virtual world, see Fig. 3. Alexiadis et al. [22] discuss in detail the reconstruction module's pipeline for capturing and 3D reconstruction of replicants.



Fig. 3 REVERIE virtual hangout scenario session with replicants

3.4 Interaction with the autonomous agent

REVERIE emphasizes non-verbal, social and emotional perception, autonomous interaction and behavioural recognition features [23]. The system can capture the user facial expression, gaze and voice, and interacts with them through an ECA's body, face gestures and voice commands. The agent exhibits audiovisual listener feedback and takes user's feedback into account in real time. The agent pursues different dialogue strategies depending on the user's state, interprets the user's non-verbal behaviour and adapts its own behaviour accordingly. To direct the participants in the VE to areas where they should focus, REVERIE uses the "Follow-Me" module. The Follow-Me module controls all user avatars' navigation in a scene by automatically addressing them destinations according to an ECA's path and the virtual scene structure. The users' facial expressions are also captured in real time through their webcam and adapted to their character's representation allowing the analysis of the users' attention and emotional status throughout the whole experience. This is controlled by the "Human Affect Analysis Module" and allows the users that may take the role of instructors in the virtual activity to be aware of the emotions and feelings of users they need to guide during a virtual experience.

4 Use cases and field studies

This section discusses two use cases that drove the development of REVERIE [24] and two field studies that have been conducted to evaluate the following:

- The impact of REVERIE's immersive and multimodal communication features
- The cognitive accessibility of educational tasks completed with the use of the platform
- The quality of the user experience (UX)

The quality of REVERIE VHs is then reviewed under the lens of a list of VH design guidelines that should be met to aid smooth interaction, communication, collaboration and control in CVEs (see Section 5).

4.1 Use case 1

The first use case (UC1) shows how REVERIE can be used in educational environments with emphasis on social networking and learning. This is accomplished through the following:

- The integration of social networking services
- The provision of tools for creating personalized look-alike avatars
- Navigation support services for avoiding collisions
- Spatial audio adaptation techniques
- Real-time facial animation adaptation for avatar representation
- Artificial intelligence techniques for responding to the users' emotional status

UC1 consists of two educational scenarios. Scenario 1 (Sc1) is a guided tour in the virtual European Parliament followed by a debate. Students log in to REVERIE using their personal account credentials and create their own personalized avatar by accessing the RAAT tool (see Section 3.1). Then, they are transferred in the virtual parliament scene where an automated explanatory tour takes place guided by an autonomous agent (see Section 3.4). When the tour is over, a virtual debate takes place and each student presents their view on a topic (e.g. multiculturalism) which is streamed as video over the internet and is being rendered in a 3D virtual projector in the parliament scene. Students rate their preferred presentations using the rating widget (at a 5-point scale) and the results are shown to everyone through an overlay list display.

Scenario 2 (Sc2) is a search-and-find game in a 3D virtual gallery. Students log in to REVERIE, create their own avatar as in Sc1 and join a virtual gallery. Each student is assigned with a card containing information about an object in the virtual gallery based on which they have to locate this object by exploring the virtual gallery on their own. Then, they have to give a presentation about the object they have found. As in Sc1, students rate their preferred presentations.

In both scenarios, communication between users in the VE is multimodal supporting the use of spatial audio streaming and avatar gestures.

4.2 UC1 field trial

UC1 field trial was conducted in a lab at Queen Mary University, London, with a setup resembling an actual classroom environment, see Fig. 4. In particular, a series of desks

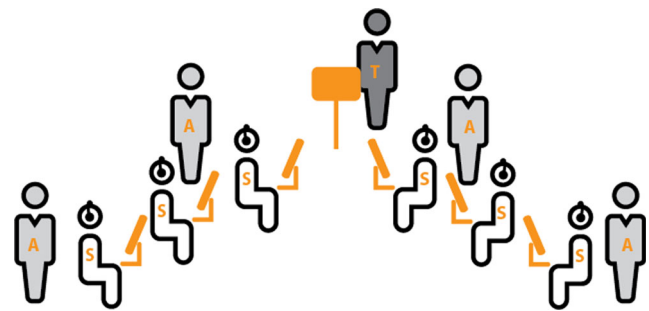


Fig. 4 REVERIE UC1 field trial setting. *S* student, *T* teacher, *A* assistant

was put in a rectangular shape, dividing the space into three areas, two sides allocated to students (represented with an *S* in Fig. 4), and the top to the teacher(s) (represented with a *T* in Fig. 4) associated with each group. Four researchers were present in the lab (represented with an *A* in Fig. 4) to record each session and to provide the necessary support (technical and logistical) for the successful completion of each session. Participants used high specification desktop computers, standard wireless mouse and a Bluetooth headset attached to a computer to complete the assigned educational tasks. Web cameras were attached to the computers to enable multimodal communication (e.g. head nods) and to detect the students' attention. The computers were connected to a LAN (local area network) with access to the WWW.

In total, 52 participants took part in this study, six of which were used in an initial pilot to ensure that the main study would run smoothly. The remaining 46 participants (students and teachers) were assigned randomly in the study conditions.

Participants included a mix of male and female students between 11 and 18 years old. All participants had a variety of familiarity with video games and social networking media portals (e.g. Facebook, Twitter). Participants were administered in groups with a maximum of six members each and they were given a short training session at the beginning of the study to familiarize with the use of REVERIE. Participants were also given a printed GUI map for REVERIE in case they would still not feel comfortable with its use after the training session. After this training, the teachers were asked to follow the lesson plan that consisted of the following:

- A *starter activity*, which was a group discussion designed to introduce and give students a chance to reflect on the topic of the educational activities
- An *individual activity*, where the students got the opportunity to present their views (with arguments for and/or against) on the topic of the educational activities using REVERIE and get feedback

- A *group activity*, where the students had to work closely with their classmates to come up with an answer on the topic of the educational activity using REVERIE VE. Each session was followed by an interview of the students discussing their experience using REVERIE.

4.3 Use case 2

The second use case (UC2) of REVERIE features a novel 3D tele-immersion system that can stream in real-time 3D meshes (including high-resolution scans of real users) which can be fully integrated into virtual worlds. Participants in a group of four had to enter a “virtual hangout” and they had to chat and collaborate in the CVE to complete the assigned task. Participants were represented in the VE with replicants who were captured with three Kinects. This was done to investigate the impact of the level of realism to the quality of the UX and task completion. The user represented by the replicant was given a step-by-step manual on how to create two objects using Lego Mega Blocks. Using verbal and non-verbal means of communication, the replicant had to show to the rest of the group the objects he/she created using the Lego Mega Blocks. The rest of the group had to replicate the shapes on a notepad using words to describe their various features (e.g. color and shape). The user represented by the replicant commented on the accuracy of the drawings and the whole process was repeated with the second object.

4.4 UC2 field trial

The evaluation of the scenario was carried out in the laboratory of Dublin City University. Thirty-one participants were recruited internally for the study with a variety of computing and educational background and with mixed gender and age. The following three dimensions of UX were deemed as important:

- *The usability of the UC2 prototype*: the degree in which participants are able to complete assigned tasks with effectiveness, efficiency and satisfaction
- *The user engagement*: the extent by the user’s experience makes the entertainment prototypes desirable to use for longer and more frequently
- *The user acceptance*: the degree by which the prototypes can handle the tasks for which they were designed [25]

Field study user experience feedback of both UCs was collected, by video recording user testing sessions, asking the users to complete questionnaires and conducting interviews after each session. Video recordings included a screengrab of individual users’ monitors documenting their actions using REVERIE, plus recording the room

that the activity took place (see Fig. 4). The later recordings provided data related to user communication outside the system and to cases where human intervention was required by human helpers attending the session. In addition, data included communication logs between all users using REVERIE communication tools. Interviews have been video recorded and transcribed detailing user reaction to a set of question related to the user impression about REVERIE’s user interface, mode of communication (text vs audio), user feeling of immersion, VH representation, user control over the activity, navigation and gamification. This resulted to a rich qualitative dataset. All recorded data was analysed by expert usability evaluators following formal methods (e.g. analyse the sense of participants’ presence in the virtual world by measuring their attention, emotional engagement and overall sentiment) [26] aiming to extract repetitive patterns in user response that could lead to quantifying user response and forming more generalizable remarks. The evaluation of the data collected from the field trials is described in Section 5 below demonstrating the results on the quality of REVERIE VHs.

5 Evaluation of REVERIE VHs

This section reviews how well REVERIE VH addresses the VH framework of design guidelines that warrant smooth social communication and interaction in CVEs and in essence support immersion. Those VH design guidelines derived from previous empirical research which is outlined in Section 5.1 below and are listed in columns 1–2 of Table 1. The data used for this analysis stems from REVERIE tools used to implement UC1 and UC2 and from reviewing the quantitative and qualitative data collected from the field trials (see Section 4). Table 1 columns 3–6 depict which design guidelines are met by the three different REVERIE VH representation artefacts (avatars, puppeted avatars and replicants) and by the autonomous agent (shown as ECA).

The following sections provide information about the methodology based on which the VH framework of design guidelines derived and the justification of how each design guideline is met by the REVERIE VHs.

5.1 Elicitation of VH design guidelines

The derivation of design guidelines that ensures smooth interaction in CVES was the output of a user-centred, iterative, multiphased approach [12] involving overall 60 users. The novel aspects of this approach are that [12]

- It uses a “real-world” application as a case study

The advantage of using a real-world application is that problems arising in such a situation can determine the success or the failure of the system according to real user and application needs

- It breaks the problem into a series of phases of increasing sophistication

Increased sophistication is achieved by incremental upsurge of user populace, use of more “mature” technologies and conducting ethnographic studies of face-to-face user actions to determine requirements that CVEs should support. Increased sophistication helps overcoming the difficulty of isolating the vast amount of factors involved in the situation in generating the required results and evaluating their validity.

- It follows a rigorous method of analysing rich qualitative data

The approach organizes and manages rich qualitative data and enables the extraction of quantitative values, which helps deriving design guidelines and technology requirements regarding the use of VHs in CVEs for learning [27].

The users that took part at this study have been engaged in an educational activity learning the rules of a game, the ancient Egyptian game (Senet), and finally playing the game within a bespoke CVE [28] using Deva Large-Scale Distributed Virtual Reality System [29]. The rules of the game have been provided by a user that took the role of an instructor (played by a researcher) who simulated the preferred behaviour of an autonomous expert agent. The study provides an understanding of interactivity and social communication issues that arise in a collaborative environment and it creates a set of design guidelines related to the CVE environment, objects contained in the CVE and VH features, behaviours and controls. In this paper, only the guidelines related to the VH are considered. Although the context of the study was learning, the design guidelines are generic and apply to a wide range of CVEs.

5.2 Aesthetically pleasing, realistic representation

This section assesses how REVERIE VHs addressed the DG related to the aesthetics of the VH representation in VEs.

- *DG1: VHs should support realistic or aesthetically pleasing representation of the user*

UC2 stretched the importance of high-quality VH representation in VE to add realism to the activity. Users stated that no realistic appearance of all types of avatars was distracting. Specifically for avatars and puppeted avatars, users identified VH features such as unrealistic tone of skin, emotionless facial expressions, bad lip synchronization and lack of non-verbal gestures (randomized gestures not realistic) particularly distracting.

Replicants increased the level of satisfaction of user representation in the VE as they provided their actual representation/clone of their body in VR. A representative quote follows: “The moments that you could actually see the replicant, well the quality was very good and looked real”. However, users felt odd in cases where only half of their body was represented in the VE or parts of their body were disappearing due to lagging.

5.3 Identity

This section reviews how REVERIE VHs addressed DGs related to representing the user’s identity in VEs.

- *DG2: VHs should support unique representation*

UC1 showed that the REVERIE RAAT (see Section 3.1) and avatar puppeting satisfy the requirement for unique VH representation which is classed by users as very important for two reasons: it closely matches users’ personality, and it helps in user identification/recognition by others in the VE. The users expressed the requirement for a bigger list of clothes and accessories to closely and more accurately represent the user’s natural appearance and convey their identity, personality and uniqueness. They stated that although such accessories do not seriously contribute to the main activities in VR, they do improve the realism of the task.

Replicants match meticulously the requirement for accurate representation of users in the VE as they provide the exact reconstruction of the users.

- *DG3: VHs should convey the user’s role in the CVE (e.g. student, teacher, other)*

This DG was fully met by REVERIE by providing models/emblems via the RAAT to represent users with specific roles.

- *DG4: VHs should support customizable behaviour*

Puppeted avatar and replicant behaviour is driven by the user. Editing of agents’ behaviour/discourse and verbal or non-verbal response is not supported by REVERIE.

5.4 Users focus of attention

This section looks how REVERIE VHs addressed DGs related to the user’s involvement in activities that take place in the VE and their level of engagement.

- *DG5: VHs should convey the user’s viewpoint*
- *DG6: an active participant needs to be identified even when their VH is out of other users’ viewpoints*

In both DG5 and DG6, the avatars’ positioning and direction of gaze indicate the user’s focus of attention in the VE. The Human Affect Analysis Module (see

Section 3.4) provides information about a user's emotional status (only in UC1). Based on this module, the user can indirectly conclude who is paying attention.

- *DG7: user viewpoints should be easily directed to see an active participant or a speaker even when they are out of other users' viewpoint*
- *DG8: a tool should be provided for users to lock onto the active VH and follow it automatically*

UC1 showed that DG7 and DG8 are met by REVERIE navigation system due to the Follow-Me module (see Section 3.4). This module affects navigation of a group of participants by a system-controlled autonomous agent that directs user attention where it should be focused. However, users found losing control of their viewpoint intrusive and they expressed discomfort with the way this happened.

5.5 Communication and turn taking

This section evaluates how REVERIE VHs addressed DGs related to enhancing discourse in CVEs.

- *DG9: a VH should be easily associated with its communication*
- *DG10: the speaker needs to be identified even when their VH is out of other users' viewpoints*

UC1 showed that REVERIE addresses DG9 and DG10 by means of the lip synchronization and Webcam Face Puppeting module. The latter allows users to control the movements of their avatar's face through direct mapping of the character's face mesh geometry to a number of tracked feature points on the user's face. When the users are distant or hidden outside one's viewpoint, REVERIE provides a list displaying the speaker and the users that request turn to talk.

- *DG11: VHs should convey the user intention to take turn or offering a turn even when not being in other users' viewpoints*

UC1 shows that REVERIE supports a turn-taking protocol and meets DG11. This is achieved by providing a tool that shows who is talking and who wants to take turn in two ways: by pressing a button that makes the user's VH hand rise to claim a turn and by showing an indication in a list that displays the speaker and who wants to take turn. UC1 and UC2 showed that puppeted avatars and replicants convey the user intention to discourse as they closely map user facial expressions.

5.6 Private communication and interaction

This section studies how REVERIE VHs addressed DGs related to private communication and interaction in CVEs.

- *DG12: private communication and interaction should be supported*
- *DG13: VHs should show when the user is involved in private communication and whether or not others could join in.*

Regarding DG12 and DG13, UC1 showed that private communication and interaction is not supported by REVERIE. However, user testing revealed the importance of extending the platform to meet this requirement. In field trials of UC1, private communication would support teams to talk/interact with each other before taking part in a debate or a public presentation. Teachers taking part in the study mentioned that private communication would benefit educational purposes, as it would allow them to deal with group or individual user questions.

5.7 User status

This section looks how the affective features of REVERIE VHs addressed DGs related to the user status of interaction with others or objects in the VE and state of mind.

- *DG14: VHs should reveal the user's action point*
- *DG15: users need to be provided with real-time cues about their own actions*

DG14 and DG15 are about associating object manipulation with the VHs performing those actions. REVERIE VHs are restricted in revealing information about the VHs walking, looking at certain directions or being seated. The animations they support are restricted to random verbal movements and not to object manipulation in the VE. In contrast, puppeted avatars fulfill those guidelines as they are capable of reconstructing an animation in the VE that copies the user action from real life. Replicants fully address those design guidelines as they provide an exact 3D reconstruction of the user and possibly objects that the user may be carrying/manipulating (including both 3D geometry and texture) in real time. However, puppeted avatars' and replicants' movement in the VE is restricted to a small physical space within the range that can be covered by the Kinect. Supporting different viewpoints in the VE allows users to see their own human representation performing an action in the VE.

- *DG16: VHs should convey explicitly the user's process of activity and state of mind.*

Two points are covered by this DG: capturing the users' process of activity, such as starting and completing moving/manipulating an object or walking or navigating to reach a point; and state of mind, meaning engagement and focus of attention. REVERIE avatars convey process of movement (reaching a place), while

puppeted avatars and replicants fully convey a process of activity. The Gaze Direction User Engagement component of the REVERIE Human Affect Analysis Module in UC1 allows users with the role of a teacher to monitor if the student users are focused on the activity (if they look at the screen) and their emotional state (happy, neutral or unhappy).

5.8 Control

This section looks how DGs related to control in an educational CVE or any other environment where similar conditions of practical management apply are addressed by REVERIE VH representation. In this design guidelines group, by expert we refer to users with the role of a teacher/instructor and by novice we refer to users that learn within the VE.

- *DG17: the expert should be in control of novice user behaviour*

DG17 implies the existence of tools that allow monitoring user behaviour and being able to intervene to aid user interaction in the VE. Such a design solution would be beneficial in any pedagogical environment. The facial puppeting capabilities provided by REVERIE meet DG17 as well, as they inform a user with the role of a teacher/instructor about student users' engagement via the Human Affect Analysis Module based on which teachers can intervene to attract students' attention.

- *DG18: the expert should have control over an individual user's viewpoint*

The Human Affect Analysis Module informs expert users about other users' emotional status. This helps an expert user to change behaviour in order to attract disengaged users. The Human Affect Analysis Module and the "Follow-Me" module (see Section 3.4) allow the autonomous agent to attract the user attention by approaching users that appear disengaged and start clapping in front of them to attract their attention and guiding users to an area of interest. Such solution has been characterized as very intrusive by the student users in UC1. However, teacher users stated that for educational purposes, this feature is necessary.

- *DG19: the expert should be in control of the communication tools*

REVERIE partially meets DG19 as controlling students' communication is restricted to muting their VH. They do not have any control over other VHs though.

- *DG20: the expert should be able to take control of objects in the CVE*

DG20 is about providing tools that allow expert users being in control of objects contained in the VE and

other VHs' behaviour. None of those requirements are currently met by REVERIE.

- *DG21: the expert should be aware of and have control over private communication of novice users*
- *DG22: the expert should be aware of and have control over private interactions of novice users*

Design guidelines 21 and 22 are not met by REVERIE as the system does not support private communication or interaction. Satisfying those requirements would be beneficial in any pedagogical environment.

- *DG23: the expert should have an episodic memory of novice user mistakes*

DG23 stretches the need for a tool that keeps history of frequent mistakes. This implies providing a tool for setting a series of actions, a set of properties (right/wrong) and keeping tracks of a user's progress of task completion in a VE. Satisfying DG23 would increase expert's promptness in assisting novice users. REVERIE partially meets DG23 by providing a tool that allows recording user actions in the VE that could be viewed by an expert user in order to provide feedback and guide other users. However, this is difficult to happen in real time and for a large number of users.

6 Discussion and directions of future work

In this section, we discuss the results of the evaluation of REVERIE VH representation based on how effectively they addressed:

- The needs of the use cases that have been created to evaluate REVERIE's technological features in real world scenarios (see Section 4)
- The user-centred VH DGs outlined in Table 1.

The results are discussed according to a list of features that VH representation should fulfill in virtual platforms that enable remote synchronous human interchange to successfully support smooth interaction, communication and collaboration, with prim focus to REVERIE. Also, this discussion leads to general remarks of directions for future development of REVERIE which are directly applicable to the CVEs that have been discussed in the state of the art of VH representation section (see Section 2). The discussion of the list of VH features is grouped as follows:

- *Aesthetically pleasing and realistic representation* is important in a VE in order to aid realism in the activity
REVERIE replicants fully support realistic representation of users in a VE, while puppeted avatars support it close enough. To fully address this requirement, direction for future work should focus towards representing

more realistically real-life user facial expressions and performance of puppeted avatars and perfecting real-life image reconstruction of replicants.

- *Identity* is important to effectively identify user roles and represent user personalities

REVERIE replicants meet meticulously the need for accurate representation of users in the VE. Avatar puppeting that maps the image of the user to the avatar face helps closely match the natural user appearance, while the REVERIE RAAT allows the customization of the user avatars to closely and accurately match the user natural appearance. The RAAT tool could be extended with a greater list of actors, clothes, hair styles and accessories to match user requirement for a more personalized representation.

- *User focus of attention and status of activity* are essential prerequisites in initiating and following up an activity on the basis of associating VHs with actions they are engaged in a VE.

The REVERIE avatars' positioning in the VE indicates the users' focus of attention, while replicants and puppeted avatars adequately reveal the action which is performed by a user in the VE. However, the field trials indicated that more work needs to be done in order to improve the quality of both sets of VH representation replicants and puppeted avatars to avoid breaking of the models and motion resulting in a non-realistic representation of users and user actions.

The REVERIE Human Affect Analysis Module (see Section 3.4) indicates users' focus of attention and addresses problems of constrained sense of actions performed in a CVEs which is imposed in general by the restricted human visual field and spatial audio in VR. However, the field trials showed that although the actual way the autonomous agent is designed to attract and refocus user attention in REVERIE (see DG18) was appreciated by teachers that need tools to enforce control in a pedagogical environment, it was generally rather intrusive for the rest of the users.

- *Communication and turn taking* is supported adequately in REVERIE by lip synchronization, avatar puppeting and relevant animations that identify the speaker expressing willingness to take turn, for example raising your avatar's hand to express interest to take turn.
- *Private communication and interaction* is not supported by REVERIE. UC1 field trials revealed the importance of extending the system to meet this requirement particularly to support pedagogical purposes that deal with assisting novice users to become more active.
- *Control* implies the need to record and take control of other users' communication and actions in the VE in order to effectively assist the activity.

Such requirement is particularly valuable in pedagogical environments where control over trainees needs to be applied to effectively assist educational requirements. This is met in REVERIE through the services provided by the Follow-Me module, the Human Affect Analysis Module and the design of the graphical user interface that allows the teacher user to control who can be heard or not in the CVE. Control also implies the need of keeping track of actions/mistakes and being able to intervene and assist the activity accordingly. This implies that the system provides a tool for setting a series of actions, a set of behaviour (right/wrong) and to track the user's progress of activities in a VE. Satisfying such a requirement would increase the expert's control over novice users' progress and comprehension and it would enhance the teacher's promptness in assisting them. This would be particularly beneficial in any pedagogical environment or any VE where users need to follow specific tasks and routines. REVERIE partially meets this requirement by providing a tool that allows recording user actions in VE that could be viewed by an expert user in order to provide feedback and guide other users. This solution might be effective for feedback following an activity, not in real time, and for a small number of users taking part. Otherwise, automation of the process is required.

7 Conclusions

The evaluation of VH representation in the REVERIE environment showed that it partially satisfies the design guidelines for realistic human representation. We believe the next steps in developing CVEs in the future should cater for all main features REVERIE currently provides or are recommended in Section 6 of discussion for further improvements of the platform. In summary, these sets of features should include the following:

- Successful indication of user focus of attention and status of activity
- Effective support of turn taking
- Detection of user status of engagement in the interactions
- Reaction in various ways by showing appropriate behaviors (gestures, gaze, speeches) in response
- Tools of control over user actions

Further VH features relating to smooth interaction, communication and collaboration in CVEs should be considered for development towards the following:

- Improving 3D reconstruction techniques for the creation of realistic VH representation

- Integrating AI tools for analysing user behaviour and engagement that will grant control to expert users when required
- Recording and being in control of communication and user actions
- Supporting private communication and interaction

In this paper, we discussed the importance of human representation in VEs to foster communication, interaction and collaboration, and as a result aid the feeling of presence and immersion in CVEs. We demonstrated a list of features that VEs should encompass in order to become more affective and support immersion in a VE. Those can be highlighted to realistic lookalike representation and behaviour, integration of communication control tools and tools for providing feedback on user's attention. For these features to be met, certain technological breakthroughs will have to be made. Realistic VE representation to reach human eye level of detail requires technological advance in compression algorithms, available internet bandwidth and increasing the resolution of 3D capturing devices. Successful puppeting of VE requires a hybrid multimodal control scheme which integrates optical systems (e.g. a Kinect device) with wearable sensors such as wireless/wearable inertial measurement unit (WIMU) which has been extensively used in REVERIE [30]. Such hybrid system could provide accurate information for both positioning and posture of the human user which should translate to better puppeting of their avatar (or replicant). Similarly, a system which fuses optical and wearable biometric sensors could also provide finer-grained information about the user's emotional state to the artificial intelligence system and improve the VE affective response. Currently, wearable sensors are big and cumbersome and are likely to be considered intrusive by most users. However, in the future, these sensors are expected to become a seamless part of human clothing so as to help in simulating reality and advancing intuitive human interaction in CVEs.

Acknowledgements The research that led to this paper was supported in part by the European Commission under the Contract FP7-ICT-287723 REVERIE.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

1. Kahneman D (1973) Attention and effort. Prentice Hall Inc, NJ
2. Goodwin C (1986) Gestures as a resource for the organisation of mutual orientation. *Semiotica* 62(1/2):29–49
3. Becker B, Mark G (1998) Social conversation in collaborative virtual environments. In: Snowdon D, Churchill E (eds) *Proceedings of the collaborative virtual environments (CVE'98)*. University of Manchester, pp 47–56
4. Ekman P, Friesen W (1978) Facial action coding system. Consulting Psychologists Press, Palo Alto
5. Sacks H (1992) Lectures on conversation. Blackwell, Cambridge
6. Argyle M (1988) Bodily communication, 2nd edn. Methuen, London
7. Benford SD, Bowers JM, Fahlén LE, Greenhalgh CM, Snowdon DN (1995) User embodiment in collaborative virtual environments. In: *Proceedings of the ACM conference on human factors in computing systems (CHI'95)*. ACM/SIGCHI, Denver, pp 242–249
8. Kileni K, Groten R, Slater M (2012) The sense of embodiment in virtual reality. *Presence: Teleoperators and Virt Environ* 21:373–387
9. Slater M, Usoh M (1993) Presence in immersive virtual environments. In: *Proceedings of the IEEE conference—virtual reality annual international symposium (VRAIS'93)*, IEEE neural networks council. IEEE Computer Society, Seattle, pp 90–96
10. http://cordis.europa.eu/project/rcn/100323_en.html. Accessed 12 July 2017
11. Economou D (2001) The role of virtual actors in collaborative virtual environments for learning. (PhD Thesis) Department of Computing and Mathematics, Manchester Metropolitan University, Manchester
12. Economou D, Pettifer SR (2005) Towards a user-centred method for studying CVEs for learning. In: *Developing future interactive systems*. Idea Group Publishing, Hershey, pp 269–301. ISBN 1591404126
13. <http://secondlife.com/>. Accessed 12 July 2017
14. http://opensimulator.org/wiki/Main_Page. Accessed 12 July 2017
15. <https://www.activeworlds.com/>. Accessed 12 July 2017
16. <https://www.sansar.com/>. Accessed 12 July 2017
17. <https://www.highfidelity.io/>. Accessed 12 July 2017
18. Fririksson FA, Kristjánsson HS, Sigursson DA, Thue D, Vilhjálmsdóttir HH (2016) Become your avatar: fast skeletal reconstruction from sparse data for fully-tracked VR. In: *Proceedings of the 26th international conference on artificial reality and telexistence and the 21st eurographics symposium on virtual environments: posters and demos*. Eurographics Assoc. Arkansas, Little Rock
19. Fechteler P, Hilsmann A, Eisert P, Broeck SV, Stevens C, Wall J, Sanna M, Mauro DA, Kuijk F, Mekuria R, Cesar P, Monaghan D, O'Connor NE, Daras P, Alexiadis D, Zahariadis T (2013) A framework for realistic 3D tele-immersion. In: *MIRAGE '13 Proceedings of the 6th international conference on computer vision / computer graphics collaboration techniques and applications*, Article No. 12. ACM, New York, pp 1–8
20. Apostolakis KC, Daras P (2013) RAAT—the reverie avatar authoring tool. In: *18th International conference on digital signal processing (DSP)*. IEEE
21. Apostolakis KC, Daras P (2015) Natural user interfaces for virtual character full body and facial animation in immersive virtual worlds. In: *International conference on augmented and virtual reality*. Springer International Publishing, pp 371–383
22. Alexiadis DS, Zarpalas D, Daras P (2013) Real-time, realistic, full 3-D reconstruction of moving humans from multiple Kinect streams. *IEEE T Multimed* 15(2):339–358
23. Kuijk F, Apostolakis KC, Daras P, Ravenet B, Wei H, Monaghan DS (2015) Autonomous agents and avatars in REVERIE's virtual environment. In: *International conference on 3D Web Technology*. Heraklion
24. Wall J, Izquierdo E, Argyriou L, Monaghan DS, O'Connor NE, Poulakos S, Mekuria R (2014) REVERIE: natural human

- interaction in virtual immersive environments. In: 2014 IEEE International conference image processing (ICIP). IEEE, pp 2165–2167
25. Davis FD, Venkatesh V (2004) Toward preprototype user acceptance testing of new information systems: implications for software project management. *IEEE Trans Eng Manag* 51(1):31–46
26. Silverman D (2006) Interpreting qualitative data: methods for analyzing talk, text and interaction. Sage
27. Economou D, Mitchell W, Boyle T (2000) Requirements elicitation for virtual actors in collaborative learning environments. *Comput Educ* 34(3–4):225–239
28. Economou D, Mitchell WL, Pettifer SR (2002) Problem driven CVE technology development. *J Comput Netw Appl (JCNA) Elsevier Science* 25(3):1–20
29. Pettifer SR, Cook J, Marsh J, West A (2000) Deva 3: architecture for a large-scale distributed virtual reality system. In: *Proceedings of ACM symposium in virtual reality software and technology*. ACM Press, pp 33–39
30. Valtazanos A, Arvind DK, Ramamoorthy S (2013) Using wearable inertial sensors for posture and position tracking in unconstrained environments through learned translation manifolds. In: *Proceedings of the 12th international conference on Information processing in sensor networks*. ACM, Philadelphia, pp 241–252